

TYPE-2 퍼지 추론 구동형 RBF 신경 회로망 설계 및 최적화

백진열, 김웅기, 오성권, 김현기  
수원대학교 전기공학과

Design of Radial Basis Function Neural Network Driven to TYPE-2 Fuzzy Inference and Its Optimization

Jin-Yeol Baek, Woong-Ki Kim, Sung-Kwun Oh, Hyun-Ki Kim  
Dept. of Electrical Engineering, The University of Suwon

**Abstract** - 본 논문에서는 TYPE-2 퍼지 추론 기반의 RBF 뉴럴 네트워크(TYPE-2 Radial Basis Function Neural Network: T2RBFNN)를 설계하고 PSO(Particle Swarm Optimization) 알고리즘을 이용하여 모델의 파라미터를 동정한다. 제안된 모델의 은닉층은 TYPE-2 가우시안 활성화 함수로 구성되며, 출력층은 Interval set 형태의 연결가중치를 갖는다. 여기에서 규칙 전반부 활성화함수의 중심 선택은 C-means 클러스터링 알고리즘을 이용하고, 규칙 후반부 Interval set 형태의 연결가중치 결정에는 경사 하강법(Gradient descent method)을 이용한 오류 역전파 알고리즘을 사용하여 학습한다. 또한, 최적의 모델을 설계하기 위한 학습을 및 활성화함수의 활성화 영역 결정에는 입자 군집 최적화(PSO; Particle Swarm Optimization) 알고리즘으로 동조한다. 마지막으로, 제안된 모델의 평가를 위하여 모의 데이터 집합(Synthetic dataset)을 적용하고 근사화 및 일반화 능력에 대하여 토의한다.

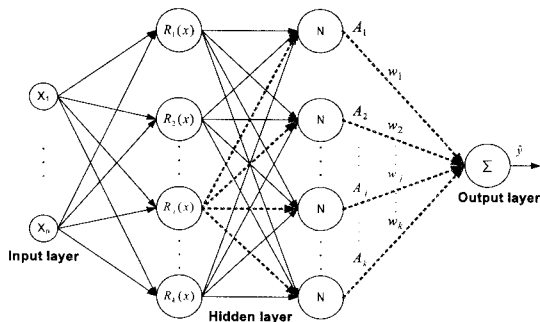
1. 서 론

퍼지 논리 시스템의 멤버십 함수와 규칙의 구조는 불확실성이 존재하는 언어적 정보 또는 수치적 데이터를 바탕으로 설계된다. 따라서 기존의 퍼지 논리 시스템의 멤버십 함수는 외부의 노이즈와 같은 불확실한 정보를 표현하는데 한계가 있다. 그러나 Zadeh의 TYPE-2 퍼지 집합[1]에 기초한 TYPE-2 퍼지 논리 시스템은 불확실한 정보를 멤버십 함수로 표현함으로써 불확실성을 효율적으로 취급할 수 있다. TYPE-2 퍼지 논리 집합은 Mendel등에 의하여 General TYPE-2 퍼지 논리 집합과 Interval TYPE-2 퍼지 논리 집합으로 구체화 되었으며[2][3], 현재 신경망과 결합을 통한 활발한 연구가 진행 중이다[4][5]. 한편, RBF 신경 회로망은 지능형 구조 중 퍼지추론 모델과 함께 가장 많이 사용되는 대표적인 모델로, 하나의 은닉층(hidden layer)만으로 구성된 네트워크 구조이다. RBF 신경 회로망은 빠른 학습시간, 일반화(generality)능력, 단순화(simplicity)의 특징으로 인해 학습 데이터를 분류하는 작업과 비선형 시스템 모델링 등에 적용되고 있다[6][7]. 따라서 본 논문에서는 TYPE-2 퍼지 추론 기반의 RBF(Radial Basis Function) 신경 회로망을 설계한다. 언어적인 불확실한 정보를 효율적으로 취급하는 TYPE-2 퍼지 추론 구동형 RBF 신경 회로망을 설계하고 수치적 실험을 통하여 제안된 모델의 근사화 및 일반화 성능에 대하여 논의한다.

2. 본 론

2.1 퍼지 추론 기반의 RBF 신경 회로망

그림 1은 퍼지 추론 기반의 RBF 신경 회로망의 구조로, 퍼지 추론의 추론부에 필요한 정규화 부분이 은닉층에 추가되어 있다. 은닉층의 활성화 함수로는 주로 가우시안 형태가 사용되며, 은닉층과 출력층 사이 실수 형태의 연결가중치는 오류역전파 알고리즘을 통하여 학습한다.



〈그림 1〉 퍼지 추론 기반의 RBF 신경 회로망의 구조

$$R_j = \exp\left(-\frac{1}{2} \left(\sum_{i=1}^n \frac{(x_i - v_{ji})^2}{(\sigma_{ji})^2}\right)\right), \hat{y}_p = \sum_{j=1}^k R_j \cdot f_j \quad (1)$$

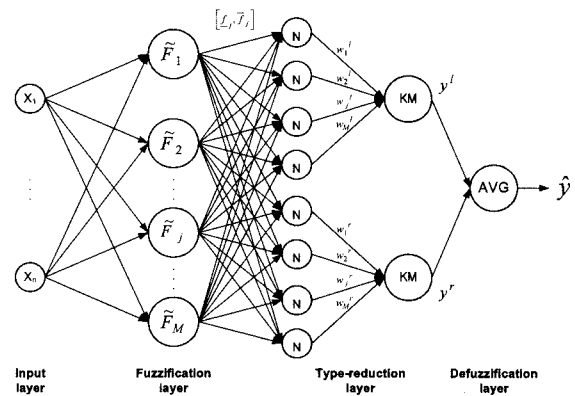
식 (1)에서 k, n은 각각 규칙과 입력의 수를 나타내고, p는 데이터를 의미한다.

2.2 TYPE-2 퍼지 추론 구동형 RBF 신경 회로망 설계

TYPE-2 퍼지 추론 구동형 RBF 신경 회로망의 구조는 그림 2와 같으며, 식 (2)와 같은 규칙으로 표현되며, Interval set 형태의 연결가중치는 식 (3)과 같다.

$$R_j = \text{If } x \text{ is } \tilde{F}_j, \text{ then } y \text{ is } \tilde{w}_j \quad (2)$$

$$\tilde{w}_j = [w_j^l, w_j^r] \quad (3)$$



〈그림 2〉 TYPE-2 퍼지 추론 구동형 RBF 신경 회로망의 구조

2.2.1 전방향 설계

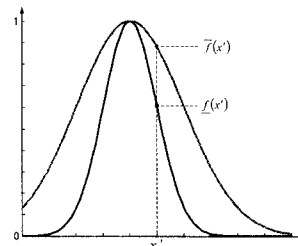
1) Input layer : 각 입력변수에 대응하는 node들로 구성되며, 입력신호를 다음 층으로 연결한다.

2) Fuzzification layer : 각 입력신호  $x_i (i=1, \dots, n)$ 에 의해 멤버십 값 즉, 적합도  $\mu_{F_j} = [f_j^l, f_j^r]$ 을 구한다.

$$F_j = \exp\left[-\frac{1}{2} \left(\sum_{i=1}^n \frac{(x_i - v_{ji})^2}{(\sigma_{ji})^2}\right)\right] = [f_j^l, f_j^r], \sigma_{ji} \in [\sigma_{ji}^l, \sigma_{ji}^r] \quad (3)$$

$$f_j^l = \exp\left[-\frac{1}{2} \left(\sum_{i=1}^n \frac{(x_i - v_{ji})^2}{(\sigma_{ji}^l)^2}\right)\right], f_j^r = \exp\left[-\frac{1}{2} \left(\sum_{i=1}^n \frac{(x_i - v_{ji})^2}{(\sigma_{ji}^r)^2}\right)\right] \quad (4)$$

식(3)은 활성화 함수로 사용한 가우시안 TYPE-2 퍼지 집합을 나타내며, n은 입력을 j는 퍼지규칙을 의미한다. 여기에서, 규칙의 수 M은 데이터들 간의 거리를 기준으로 근접한 정도를 측정하고, 이를 바탕으로 데이터를 특성별로 분류하는 HCM 클러스터링을 이용하여 결정한다. 즉, HCM에 의한 클러스터의 수와 클러스터의 중심을 제안된 모델의 TYPE-2 RBF 수와 중심 v로 사용한다. 또한 각 규칙의 활성화 영역 결정에는 최적화 알고리즘을 사용하여 동조한다. 그림 3에서는 불확실한 너비를 갖는 TYPE-2 RBF를 나타낸다.



〈그림 3〉 Uncertain standard deviation을 갖는 Interval TYPE-2 퍼지 집합

3) Type-reduction layer(Center of sets type-reduction) : 반복적인 절차를 통하여 최소 가중평균과 최대 가중평균을 계산한다.

$$y^l = \min_{f_j \in [f_j^l, f_j^r]} \left( \sum_{j=1}^M w_j^l \cdot f_j^l / \sum_{j=1}^M f_j^l \right) \quad (4)$$

$$y^r = \max_{f_j \in [\underline{f}_j, \bar{f}_j]} \left( \sum_{i=1}^M w_j^r \cdot f_j \right) \quad (5)$$

4) Defuzzification layer : 이전 층에서 얻은 모델의 최소출력과 최대출력의 평균을 취하여 모델의 최종 출력  $\hat{y}$  를 계산한다.

$$\hat{y} = (y^l + y^r) / 2 \quad (6)$$

### 2.2.2 역방향 설계

규칙의 전만부 멤버쉽 함수의 정점 결정에는 데이터들 간의 euclidean distance를 기준으로 중심을 결정하는 C-means 클러스터링 알고리즘을 사용하고, 후반부의 경우 오류 역전파(back-propagation; BP) 알고리즘을 이용하여 학습한다. 규칙 후반부 Interval set 형태의 연결가중치 동정을 위한 오차식 정의는 식 (7)로 표현되며, 전체 학습데이터에 대한 모델의 최소 오차를 얻기 위해 학습은 300회 이루어진다.

$$E_p = \frac{1}{2} (y_p - \hat{y}_p)^2 \quad (7)$$

여기서,  $E_p$ 는  $p$ 번째 데이터에 대한 모델의 오차이며,  $y_p$ 는  $p$ 번째 입력에 대한 원 출력이다.  $\hat{y}_p$ 는  $p$ 번째 입력에 대해 추론된 모델의 결과로 식 (6)에 의한다. 경사하강법에 의한 연결가중치의 학습은 식 (7)로 표현되는 오차가 최소가 되도록 식 (8)과 같이  $j$ 번째 규칙 left hand 연결가중치의 변화분  $\Delta w_j^l$ 을 이용하여 현재의 연결가중치를 갱신한다. right hand 연결가중치의 학습은 식 (9)에 의하여 수행한다.  $\eta_w^l$ 과  $\eta_w^r$ 는 학습율로 PSO 알고리즘을 이용하여 탐색한다.

$$w_j^l(new) = w_j^l(present) + \eta_w^l \cdot \Delta w_j^l \quad (8)$$

$$w_j^r(new) = w_j^r(present) + \eta_w^r \cdot \Delta w_j^r \quad (9)$$

$E_p$ 와  $w_j^l$ 사이의 관계를 사슬 법칙으로 나타내면 식 (10)과 같으며, 식 (11)로 정리된다.

$$\Delta w_j^l = - \frac{\partial E_p}{\partial w_j^l} = - \frac{\partial E_p}{\partial y} \cdot \frac{\partial y}{\partial y^l} \cdot \frac{\partial y^l}{\partial w_j^l} \quad (10)$$

$$= -(\hat{y} - y) \cdot \frac{1}{2} \cdot \frac{\bar{f}_j}{\sum_{j=1}^M \bar{f}_j} \quad (11)$$

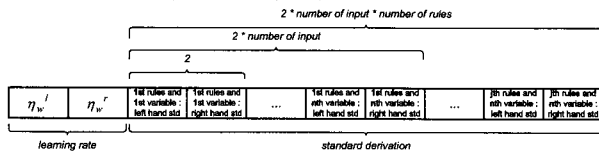
여기에서,  $\bar{f}_j$ 는 Type-reduction layer의 반복적인 절차를 통하여 결정된 것으로, 모델의 최소출력을 얻는데 선택된 각 규칙의 적합도이다. 즉,  $f_j$  또는  $\bar{f}_j$  중 하나의 값을 갖는다.  $E_p$ 와  $w_j^r$ 사이의 관계를 사슬 법칙으로 나타내면, 최종식은 식 (12)와 같다.

$$\Delta w_j^r = - \frac{\partial E_p}{\partial w_j^r} = - \frac{\partial E_p}{\partial y} \cdot \frac{\partial y}{\partial y^r} \cdot \frac{\partial y^r}{\partial w_j^r} = -(\hat{y} - y) \cdot \frac{1}{2} \cdot \frac{\bar{f}_j}{\sum_{j=1}^M \bar{f}_j} \quad (12)$$

마찬가지로, 식 (12)에서  $\bar{f}_j$ 는 반복적인 절차를 통하여 모델의 최대출력을 얻는데 선택된 각 규칙의 적합도를 나타낸다.

### 2.2.3 T2RBFNN의 최적화

제안된 모델의 연결가중치를 학습하기 위해 동정해야 하는 파라미터로 식 (8)-(9)의 학습율은 입자 군집 최적화 알고리즘(Particle Swarm Optimization)을 통하여 동정한다. 또한 활성화 함수의 최적 활성화 영역 결정을 위한 너비상수  $\sigma$ 의 탐색이 수행되며, particle의 구조는 그림 4와 같다. PSO의 초기 파라미터 값은 표 1과 같다.



〈그림 4〉 particle의 구조

〈표 1〉 PSO의 초기 파라미터들

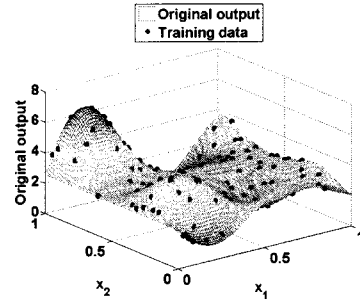
| 파라미터                            | 값                   |                      |
|---------------------------------|---------------------|----------------------|
| 세대 수                            | 60                  |                      |
| Particle                        | 50                  |                      |
| 최대 이동 속도 ( $v_{max}$ )          | 탐색공간의 20%           |                      |
| 관성 계수 [ $w_{min}$ , $w_{max}$ ] | [0.4, 0.9]          |                      |
| 가속 상수 ( $c_1$ , $c_2$ )         | 2.0, 2.0            |                      |
| 탐색 공간                           | learning rate       | [1.0e-6, 1.0-1.0e-6] |
|                                 | standard derivation |                      |

### 3. 시뮬레이션 및 결과고찰

제안된 모델의 평가에 사용된 2입력 1출력의 비선형 함수는 식 (13)으로

표현되며, 그림 5와 같다.

$$z = 1.9 \{ 1.35 + e^x e^{-x} \sin(13(x_2 - 0.6)^2) \sin(7x_1) \}, 0 \leq x_1, x_2 \leq 1 \quad (13)$$



〈그림 5〉 2입력 1출력의 비선형 함수

식 (13)의 함수로부터 그림 5에 보인 200개의 데이터를 랜덤하게 발생시켜 그 중 140개를 학습데이터로 사용하였으며, 나머지 60개의 데이터를 테스트 데이터로 사용하였으며, 제안된 모델의 근사화 성능과 일반화 성능을 평가하기 위한 지표로 식 (14)의 mean square error(MSE)를 사용하였다.

$$PI \text{ or } EPI = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \quad (14)$$

여기서 PI(Performance Index)는 학습데이터에 대한 성능지수를, EPI(Evaluated Performance Index)는 테스트 데이터에 대한 성능지수를, m은 데이터 수를 나타낸다. 제안된 모델과 기존 모델과의 성능은 표 2에 나타낸다.

〈표 2〉 T2RBFNN의 성능과 기존 모델과의 비교

| 모델              | PI       | EPI      | 모델      | PI       | EPI      | 규칙의 수 |
|-----------------|----------|----------|---------|----------|----------|-------|
| optimized RBFNN | 0.897553 | 0.771047 | T2RBFNN | 0.760902 | 0.610145 | 2     |
|                 | 0.912396 | 0.817573 |         | 0.405307 | 0.691583 | 3     |
|                 | 0.681571 | 0.783301 |         | 0.222107 | 0.436808 | 4     |
|                 | 0.484011 | 0.526575 |         | 0.145587 | 0.310189 | 5     |
|                 | 0.282569 | 0.300044 |         | 0.107565 | 0.219322 | 7     |
|                 | 0.237502 | 0.330739 |         | 0.091721 | 0.281096 | 9     |

### 4. 결 론

본 논문에서는 타입-2 퍼지 추론 구동형 RBF 신경 로망을 설계하고, 모델의 파라미터를 입자 군집 최적화 알고리즘으로 동정하였다. 활성화수와 각 규칙의 연결가중치를 interval fuzzy sets 형태로 구성한 T2RBFNN은 기존의 모델보다 근사화 및 일반화 성능에서 더 작은 오차의 보임을 표 2를 통하여 증명하였다.

### 감사의 글

이 논문은 2008년도 한국전자통신연구원의 재원으로 지원을 받아 수행된 연구임 (과제번호:2008-35MIC). 그리고 이 논문은 2007년도 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임 (KRF-2007-313-D00710).

### [참고 문헌]

- [1] L.A Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-1", Information Sciences, vol.8, pp.199-249, 1975.
- [2] J.M. Mendel, "Advances in type-2 fuzzy sets and systems", Information Sciences, vol.177, pp.84-110, 2007.
- [3] J.M. Mendel, R.I. John, and F. Lui, "Interval type-2 fuzzy logic system made simple", IEEE Trans. on Fuzzy System, vol.14, pp.808-821, Dec. 2006.
- [4] Chia-Feng Juang, Chun-Feng Lu, and Yu-Wei Tsao, "A Self-Evolving Interval Type-2 Fuzzy Neural Network for Nonlinear Systems Identification", Proceedings of the 17th World Congress, The International Federation of Automatic Control, pp.7588-7593, July 2008
- [5] Ching-Hung Lee, Jang-Lee Hong, Yu-Ching Lin, and Wei-Yu Lai, "TYPE-2 Fuzzy Neural Network Systems and Learning", International Journal of Computational Cognition, vol.1, no.4, pp.79-90, Dec 2003
- [6] Richard O. Duda, Peter E. Hart, David G. Stork, "Pattern Classification; Second Edition", John Wiley&Sons, INC., 2000.
- [7] A. Staiano, J. Tagliaferri, W. Pedrycz, "Improving RBF networks performance in regression tasks by means of a supervised fuzzy clustering Automatic structure and parameter", Neurocomputing, Vol. 69, pp. 1570-1581, 2006.